

RESEARCH ARTICLE OPEN ACCESS

Integrated Genomic and Single-Cell Analysis Reveals Heterogeneity, Prognosis, and Treatment Vulnerability in Urothelial Carcinoma

Chaozhi Tang¹ | Yifan Liu² | Tian-Long Wang³ | Jiakang Ma⁴ | Weihua Chen¹ | Xiang Liu⁵ | Jingdong Xue³ | Lin Ye¹ | Feng Yue⁶

¹Department of Urology, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China | ²Department of Urology, Tongji Hospital, School of Medicine, Tongji University, Shanghai, China | ³Department of Urology, Shanghai Eighth People's Hospital, Shanghai, China | ⁴The First Affiliated Hospital, College of Clinical Medicine of Henan University of Science and Technology, Shanghai, China | ⁵Department of Urology, Putuo People's Hospital, School of Medicine, Tongji University, Shanghai, China | ⁶Department of Urology, The Affiliated Huizhou Hospital, Guangzhou Medical University, Huizhou, Guangdong, China

Correspondence: Jingdong Xue (xuejingdong@126.com) | Lin Ye (ericyclin@tongji.edu.cn) | Feng Yue (yuefeng@gzhmu.edu.cn)

Received: 7 February 2026 | **Revised:** 21 April 2026 | **Accepted:** 6 May 2026

Guest Editor: Lin Zhang

Keywords: immunotherapy response | prognostic assessment | signature subtypes | urothelial carcinoma

ABSTRACT

At the transcriptomic level, several molecular subtyping schemes have been established to elucidate the intrinsic heterogeneity of urothelial carcinoma and to inform prognostic assessment and therapeutic guidance. However, a unified molecular classification scheme characterizing genomic alterations is lacking. Unsupervised and supervised clustering identified two distinct mutational signature subtypes. Kaplan–Meier analysis demonstrated that patients with the MUT2 subtype had a higher risk of death than those with the MUT1 subtype across multiple cohorts, including IMvigor210 (hazard ratio [HR], 1.74; 95% confidence interval [CI], 1.27–2.37; $p < 0.001$), UC-GENOME (HR, 1.54; 95% CI, 0.93–2.54; $p = 0.091$), The Cancer Genome Atlas (TCGA; HR, 1.45; 95% CI, 1.06–1.98; $p = 0.020$), MSK2022 (HR, 1.34; 95% CI, 1.10–1.64; $p = 0.004$), MSK2015 (HR, 3.43; 95% CI, 1.36–8.64; $p = 0.005$), and the Tongji cohort (HR, 4.99; 95% CI, 0.57–43.69; $p = 0.11$). Immunotherapy response rates were significantly higher in the MUT1 subtype than in the MUT2 subtype in IMvigor210 (31.8% vs. 13.1%; $p = 0.003$) and UC-GENOME (42.3% vs. 29.0%; $p = 0.022$). Consistent with these findings, single-cell analysis showed that MUT2 tumors were enriched in tumor-associated fibroblast subpopulations and had a lower abundance of immune effector cells. Overall, this genomic analysis identified two mutation-based subtypes of urothelial carcinoma associated with patient prognosis and immunotherapy response.

1 | Introduction

Bladder cancer (BC) is a prevalent and lethal malignancy worldwide [1], and its development is associated with risk factors, such as smoking, advancing age, male sex, and certain occupational exposures [2–5]. Although therapeutic options continue to evolve, the mainstays of treatment include surgery, chemotherapy, and immunotherapy [6]. Nevertheless, patient outcomes vary markedly between individuals; tumor molecular

characteristics and mutational profiles are thought to contribute to prognosis and treatment response [7, 8]. In recent years, molecular subtyping of BC has emerged as a critical factor in understanding the disease biology and developing personalized treatment strategies [9–11]. For instance, Choi et al. [12] used gene expression profiling to classify muscle-invasive BC into luminal, basal, and p53-like subtypes. The basal subtype is characterized by squamous differentiation and greater aggressiveness and is associated with shorter survival, whereas the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Copyright © 2026 Chaozhi Tang et al. *Human Mutation* published by John Wiley & Sons Ltd.

p53-like subtype exhibits significant resistance to chemotherapy. Building on these findings, Kamoun et al. [13] leveraged large-scale transcriptomic data to establish the first consensus molecular classification of muscle-invasive BC, which aids in patient risk stratification but does not incorporate treatment-related data.

Studies have indicated that mutations in genes such as *FGFR3*, *TP53*, *MDM2*, *ERCC2*, *PIK3CA*, and *RBI* are associated with tumor aggressiveness, metastatic potential, and responses to chemotherapy and immunotherapy [14–25]. In upper tract urothelial carcinoma (UC), Fujii et al. [26] delineated five subtypes based on the mutation status of *TP53*, *MDM2*, *RAS*, and *FGFR3*, revealing distinct molecular and prognostic features. Nevertheless, single-gene alterations alone are insufficient to capture the complexity of the cancer genomic landscape or fully explain tumor biology and therapeutic responses [27]. Mutational signatures—characteristic combinations of mutation types resulting from specific mutagenic processes—reflect the biological mechanisms or external influences to which a tumor has been exposed during its development [28, 29].

Compared with traditional bulk sequencing, single-cell sequencing technologies (e.g., scRNA-seq and scATAC-seq) and spatial transcriptomics enable the investigation of tumor heterogeneity and dynamic microenvironmental changes at single-cell resolution [30, 31]. These approaches overcome the limitations of signal averaging inherent in bulk methods, thereby uncovering key cellular subsets—such as cancer stem cells—and tracking their phenotypic evolution during disease progression or in response to therapy [32, 33]. Such insights are particularly valuable for elucidating mechanisms of drug resistance, as single-cell profiling can identify adaptive cellular states and their molecular signatures within resistant tumors, revealing how transcriptional reprogramming or microenvironmental interactions contribute to therapeutic escape [33–35]. Collectively, single-cell and spatial technologies provide an unprecedented level of resolution in cancer biology, enabling the characterization of cellular heterogeneity, reconstruction of dynamic regulatory networks, and identification of novel therapeutic targets.

Using an unsupervised mutational signature-based clustering approach, we identified two stable molecular subtypes of UC. In this study, we identified intrinsic mutational subtypes of UC and investigated, at single-cell resolution, the differential tumor microenvironment associated with these subtypes. We further explored their therapeutic vulnerabilities, particularly in the context of immunotherapy. In addition, we evaluated the robustness and stability of this classification model using machine learning approaches.

2 | Materials and Methods

2.1 | Data Source

Public genomic data from The Cancer Genome Atlas (TCGA) for primary BC [36], as well as immunotherapy cohorts from the IMvigor210 trial [37, 38] and Project UC-GENOME [39]

for metastatic UC of the bladder, were used in this study. Additionally, tumor and matched normal DNA libraries were sequenced using the Memorial Sloan Kettering–Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) targeted sequencing platform in 2015 [40] and 2022 [41]; these datasets are referred to as the MSK2015 and MSK2022 cohorts, respectively. Whole-exome sequencing (WES) data from patients with upper tract UC at Shanghai East Hospital (the Tongji cohort) were included as an independent, retrospectively sequenced cohort outside a clinical trial setting. This study was reviewed and approved by the Ethical Committee of Shanghai East Hospital (No. 2018YS-04) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients.

2.2 | Molecular Profiling

DNA libraries were prepared according to Illumina-recommended protocols. A total of 1 μ g of DNA was fragmented into approximately 200–300 bp fragments using a Covaris S220 ultrasonicator (Woburn, MA, United States). The DNA fragments were end-repaired and ligated to adaptors. Purified products from 10 PCR cycles were incubated with capture probes at 65°C for 24 h for exome capture. The libraries were then subjected to massively parallel sequencing using an Xten system (Illumina, Shanghai, China). Trimmomatic [42] was used for data trimming, and the Burrows–Wheeler Aligner [43] was used to align reads to the human reference genome hg38. Variant calling was performed using SAMtools mpileup output pipelines [44], and variant annotation was conducted using ANNOVAR [45], GATK [46], and VEP [47].

2.3 | Mutational Signatures and Subtyping

Mutation datasets obtained through either WES or targeted gene panel sequencing, comprising single-nucleotide polymorphisms mapped to reference genomes, were analyzed using the R package *SomaticSignatures* [48] (R Version 4.0.1, R Foundation for Statistical Computing, Vienna, Austria) to identify mutational signatures from the Catalog of Somatic Mutations in Cancer (COSMIC) v3 database [49]. Cosine similarities (CSs) were calculated across 79 single-base substitution (SBS) signatures (Supporting Information 1: Methods S1). TCGA cohort was used as the training set. Mutation signatures with CS < 0.24 present in fewer than 10% of the samples were excluded, resulting in 46 retained signatures. To determine the optimal classification, non-negative matrix factorization was applied to the CS matrix of mutational signatures (Supporting Information 1: Methods S2, Supporting Information 2: Figure S1, Supporting Information 3: Figure S2, Supporting Information 4: Figure S3, Supporting Information 5: Figure S4, and Supporting Information 6: Figure S5). Two clusters were selected as the optimal solution based on the evaluation of clustering stability. A prediction analysis for a microarray classifier was subsequently trained on TCGA cohort and used to assign mutational subtypes to external cohorts using the R package *pamr*. Calculation of the risk score is described in Supporting Information 1: Methods S3, Supporting Information 7: Figure S6, and Supporting Information 17:

Table S1. Briefly, “risk” refers to a mutation-based prognostic score developed in the MSK2022 cohort and dichotomized at a survival-optimized threshold of 1 (low risk < 1; high risk ≥ 1).

2.4 | Analysis for Single-Cell Sequencing Data

Single-cell datasets GSE169379 (25 BC samples), PRJNA662018 (eight BC cases), and GSE211388 (10 human bladder tumors) were integrated. Seurat objects [50] were generated for each dataset with quality control filters requiring a minimum of 200 features and no cells with fewer than three detected genes. Cells with fewer than 200 or more than 5000 detected features (nFeature_RNA) or with > 20% mitochondrial gene content were excluded. The combined dataset was normalized using *NormalizeData* and scaled using *ScaleData*. Variable features were selected using the “vst” method, with the Top 2000 features retained. Principal component analysis (PCA) was performed, followed by batch-effect correction using Harmony based on the “orig.ident” variable. Clustering was conducted on Harmony-corrected data with a resolution parameter of 0.4. UMAP and t-SNE were used for dimensionality reduction and visualization.

2.5 | Gene Set Variation Analysis (GSVA)

GSVA is a nonparametric, unsupervised method that estimates the enrichment of predefined gene sets at the individual sample level, generating scores that reflect pathway or biological process activity [51]. In bulk tumor sequencing, GSVA enables the identification of pathway-level expression patterns and biological signatures that may not be apparent from single-gene analyses, facilitating the investigation of tumor heterogeneity and its associations with clinical outcomes.

2.6 | Machine Learning Validation of Mutational Subtypes

Machine learning approaches, including Random Forest [52], LightGBM [53], XGBoost [54], CatBoost [55], AdaBoost [56], and logistic regression [57], were used to evaluate classification performance. TCGA cohort was randomly split into training (70%) and testing (30%) sets. External validation was performed using independent datasets such as IMvigor210 [37] to assess model robustness and generalizability (Supporting Information 1: Methods S4).

2.7 | Statistical Analysis

Continuous variables were compared using the Wilcoxon rank-sum test. Survival outcomes were evaluated using the log-rank test, and categorical variables were compared using the chi-square test. Analyses were performed using R (Version 4.0.1; R Foundation for Statistical Computing, Vienna, Austria), with p values < 0.05 considered statistically significant. Cox proportional hazard regression was used to estimate the risk of death between MUT1 and MUT2 subtypes, adjusting for age, sex, and smoking status. Survival analyses were performed using

samples with available overall survival (OS) data, excluding cases with missing OS information.

3 | Results

3.1 | Demographic and Clinical Characteristics

Table 1 summarizes the clinical and genomic characteristics of patients in the IMvigor210 ($n = 244$), UC-GENOME ($n = 191$), TCGA ($n = 412$), MSK2022 ($n = 1278$), MSK2015 ($n = 83$), and Tongji ($n = 21$) cohorts. Among the 2229 patients, 1689 (75.8%) were men, and 536 (24.1%) were women, with a mean age of 66.0 years (interquartile range [IQR], 59.0–74.0 years). In terms of smoking status, 294 (13.2%) patients were current smokers, 1157 (51.9%) were former smokers, and 709 (31.8%) were never smokers. A total of 871 (39.1%) OS events were observed. The mean follow-up time was 31.4 months (IQR, 11.4–43.4 months).

3.2 | MUT2 Subtype Predicts Unfavorable Prognosis

Among the 244 patients with metastatic UC treated with immunotherapy, 150 (61.5%) were classified as MUT1 and 94 (38.5%) as MUT2 (Table 1). The median OS for patients with the MUT2 subtype was 7.5 months, which was shorter than both the overall cohort median (9.52 months) and the 15-month median OS observed for the MUT1 subtype (hazard ratio [HR], 1.74; 95% confidence interval [CI], 1.27–2.37; $p < 0.001$) (Figure 1).

Consistent trends were observed across other cohorts, including UC-GENOME (HR, 1.54; 95% CI, 0.93–2.54; $p = 0.091$) and TCGA (HR, 1.45; 95% CI, 1.06–1.98; $p = 0.020$). The association between MUT2 subtype and shorter survival was also evident in MSK2022 (HR, 1.34; 95% CI, 1.10–1.64; $p = 0.004$) and MSK2015 (HR, 3.43; 95% CI, 1.36–8.64; $p = 0.005$). Similarly, in the Tongji cohort, MUT2 was associated with poorer survival (HR, 4.99; 95% CI, 0.57–43.69; $p = 0.11$).

3.3 | MUT1 Subtype Indicates Improved Immunotherapy Response

To explore the basis of survival differences between subtypes, we compared frequently mutated genes and clinical characteristics. No individual gene mutation showed a significant association with immunotherapy response (Supporting Information 8: Figure S7A). In contrast, mutational subtypes demonstrated clear differences: Patients with the MUT1 subtype had significantly higher response rates to immunotherapy than those with the MUT2 subtype in both IMvigor210 (31.8% vs. 13.1%; χ^2 test $p = 0.003$) and UC-GENOME (42.3% vs. 29.0%; χ^2 test $p = 0.022$) (Supporting Information 8: Figure S7B). This disparity was also reflected in tumor immune phenotypes. MUT1 tumors showed a higher proportion of inflamed phenotypes compared with MUT2 in IMvigor210 (36.6% vs. 18.8%; χ^2 test $p = 0.021$) and UC-GENOME (56.9% vs. 32.0%; χ^2 test $p = 0.007$) (Supporting Information 3: Figure S7C). Additionally, MUT1 tumors exhibited higher tumor mutational burden than MUT2 tumors (Supporting Information 8: Figure S7D,E and Supporting Information 17: Tables S2, S3, S4, S5, and

TABLE 1 | Demographic characteristics of the cohorts.

Characteristic	IMvigor210 (<i>n</i> = 244)	UC-GENOME (<i>n</i> = 191)	TCGA (<i>n</i> = 412)	MSK2022 (<i>n</i> = 1278)	MSK2015 (<i>n</i> = 83)	Tongji (<i>n</i> = 21)
Age, mean (SD)	NA	65.8 (10.1)	68.1 (10.6)	65.2 (11.2)	67.7 (10.3)	67.8 (11.7)
Sex, no. (%)						
Female	50 (20.5)	47 (24.6)	108 (26.2)	296 (23.2)	28 (33.7)	7 (33.3)
Male	194 (79.5)	143 (74.9)	304 (73.8)	980 (76.7)	54 (65.1)	14 (66.7)
NA	0 (0)	1 (0.5)	0 (0)	2 (0.1)	1 (0.2)	0 (0)
Smoker, no. (%)						
Former	145 (59.4)	112 (58.6)	198 (48.1)	650 (50.9)	44 (53.0)	8 (38.1)
Never	74 (30.3)	62 (32.5)	111 (26.9)	433 (33.9)	22 (26.5)	7 (33.3)
Current	25 (10.3)	16 (8.4)	90 (21.8)	141 (11.0)	16 (19.3)	6 (28.6)
NA	0 (0)	1 (0.5)	13 (3.2)	54 (4.2)	1 (1.2)	0 (0)
Subtype, no. (%)						
Hypermutated	1 (0.4)	2 (1.0)	1 (0.2)	7 (0.5)	1 (1.2)	1 (4.8)
FGFR3	41 (16.8)	26 (13.6)	44 (10.7)	241 (18.9)	39 (47.0)	1 (4.8)
RAS	7 (2.9)	8 (4.2)	30 (7.3)	68 (5.3)	11 (13.2)	4 (19.0)
TP53/MDM2	99 (40.6)	105 (55.0)	199 (48.3)	629 (49.2)	14 (16.9)	6 (28.6)
Triple negative	96 (39.3)	50 (26.2)	138 (33.5)	333 (26.1)	18 (21.7)	9 (42.9)
TMB, no. (%)						
< 10 Muts/Mb	213 (87.3)	76 (39.8)	315 (76.5)	672 (52.6)	61 (73.5)	19 (90.5)
≥ 10 Muts/Mb	31 (12.7)	81 (42.4)	97 (23.5)	606 (47.4)	22 (26.5)	2 (9.5)
NA	0 (0)	34 (17.8)	0 (0)	0 (0)	0 (0)	0 (0)
Risk, no. (%)						
Low	121 (49.6)	97 (50.8)	173 (42.0)	608 (47.6)	46 (55.4)	9 (42.9)
High	123 (50.4)	94 (49.2)	239 (58.0)	670 (52.4)	37 (44.6)	12 (57.1)
Overall survival, no. (%)						
Event	159 (65.2)	78 (40.8)	158 (38.3)	445 (34.8)	25 (30.1)	6 (28.6)
Censored	85 (34.8)	85 (44.5)	250 (60.7)	793 (62.1)	43 (51.8)	15 (71.4)
NA	0 (0)	28 (14.7)	4 (1.0)	40 (3.1)	15 (18.1)	0 (0)
MS subtype, no. (%)						
MUT1	150 (61.5)	63 (33.0)	161 (39.1)	833 (65.2)	37 (44.6)	12 (57.1)
MUT2	94 (38.5)	128 (67.0)	251 (60.9)	445 (34.8)	46 (55.4)	9 (42.9)

Note: Risk, mutation-based genomic risk score (Supporting Information 1: Methods S3): low risk, score < 1; high risk, score ≥ 1. Abbreviations: NA, not available; SD, standard deviation; TCGA, The Cancer Genome Atlas; TMB, tumor mutation burden.

S6), consistent with current understanding of immunotherapy responses in UC.

Subgroup analyses further supported these findings. Among patients with an Eastern Cooperative Oncology Group (ECOG) performance status of 1, MUT2 was associated with a higher risk of death than MUT1 (IMvigor210: HR, 2.47; 95% CI, 1.65–3.70; $p < 0.001$; UC-GENOME: HR, 2.96; 95% CI, 1.03–8.46; $p = 0.04$) (Supporting Information 9: Figure S8 and Supporting

Information 10: Figure S9). Similarly, in patients who received platinum chemotherapy (IMvigor210: HR, 2.02; 95% CI, 1.41–2.90; $p < 0.001$) or derived a clinical benefit from chemotherapy (UC-GENOME: HR, 8.73; 95% CI, 1.07–71.43; $p = 0.04$), the MUT2 subtype conferred a higher risk of death than the MUT1 subtype. This increased risk for the MUT2 subtype (vs. the MUT1 subtype) was also observed in the subgroup of former smokers (IMvigor210: HR, 2.00; 95% CI, 1.32–3.03; $p < 0.001$; UC-GENOME: HR, 1.79; 95% CI, 0.94–3.40; $p = 0.08$) and in the

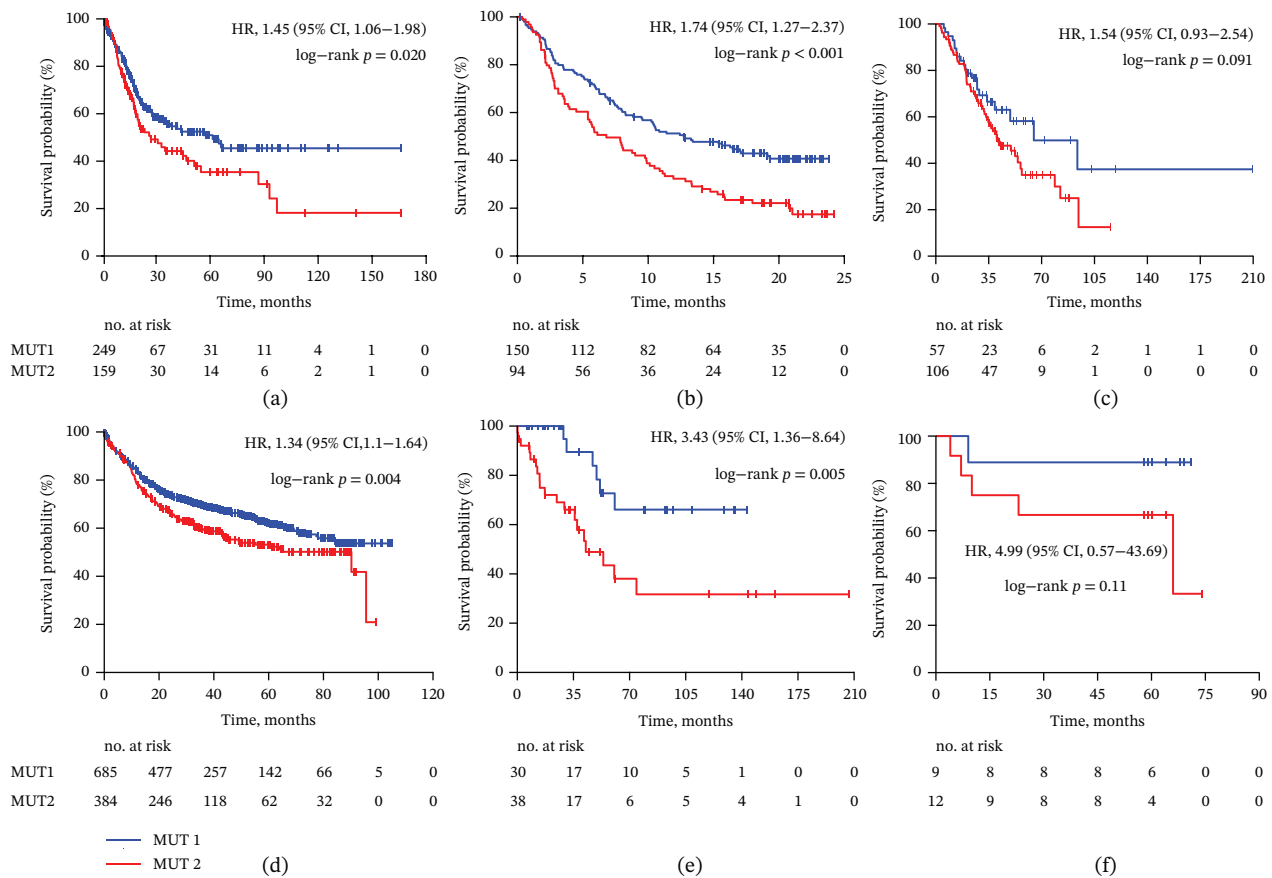


FIGURE 1 | Survival outcomes according to the mutational subtype in all retrospective cohorts. Kaplan–Meier survival analysis results showing differences in overall survival for mutational signature subtypes in the (a) TCGA, (b) IMvigor210, (c) UC-GENOME, (d) MSK2022, (e) MSK2015, and (f) Tongji self-sequencing cohorts. Hazard ratios (HRs), 95% confidence intervals (CIs), and p values are indicated on each plot, along with the number of patients at risk.

“excluded” immunophenotype subgroup (IMvigor210: HR, 1.57; 95% CI, 0.96–2.56; $p = 0.07$; UC-GENOME: HR, 2.08; 95% CI, 0.92–4.74; $p = 0.08$).

In the IMvigor210 cohort, the MUT2 subtype was associated with a higher risk of death than the MUT1 subtype in patients with liver (HR, 1.88; 95% CI, 1.09–3.24; $p = 0.02$) or visceral metastases (HR, 1.73; 95% CI, 1.09–2.74; $p = 0.02$). It was also associated with a higher risk in patients with an IC PD-L1 score of 0 at enrollment (IC0; HR, 2.05; 95% CI, 1.13–3.72; $p = 0.02$), an IC-only PD-L1 score of 0 (IC0; HR, 2.37; 95% CI, 1.28–4.29; $p = 0.01$), or a tumor cell PD-L1 score of 0 (TC0; HR, 1.96; 95% CI, 1.37–2.82; $p < 0.001$). Additionally, the MUT2 subtype was associated with significantly worse survival in the “desert” immunophenotype subgroup (HR, 2.37; 95% CI, 1.27–4.45; $p < 0.001$).

3.4 | Subgroup Analysis

Univariate analysis demonstrated that the mutational subtype remained a stable prognostic factor compared with clinical variables such as sex, age, smoking status, and tumor mutational burden. In IMvigor210, MUT2 was associated with increased risk of death compared with MUT1 (HR, 1.74; 95% CI, 1.27–2.37; $p < 0.001$). In the UC-GENOME cohort, the HR for the MUT2 subtype versus the MUT1 subtype was 1.54 (95% CI, 0.93–2.54;

$p = 0.09$). TCGA cohort showed an HR of 1.48 (95% CI, 1.10–1.98; $p = 0.009$), while the MSK2022 cohort exhibited an HR of 1.37 (95% CI, 1.12–1.69; $p = 0.003$). In the MSK2015 cohort, the HR for the MUT2 subtype versus the MUT1 subtype was 3.43 (95% CI, 1.36–8.64; $p = 0.009$) (Supporting Information 11: Figure S10). Multivariate analysis confirmed that the mutational subtype was an independent prognostic factor in the MSK2022 cohort (HR, 1.34; 95% CI, 1.07–1.67; $p = 0.011$) (Supporting Information 12: Figure S11).

3.5 | Cell Type Identification of Single-Cell Sequencing for UC

To elucidate the underlying mechanisms responsible for the disparities in prognosis and immunotherapy response between these two mutational subtypes, we performed a single-cell-level interrogation of tumor microenvironment heterogeneity. This approach allowed us to systematically characterize the differential composition, transcriptional states, and cell–cell communication networks across distinct cellular clusters—including malignant, immune, and stromal populations—within each subtype. Single-cell RNA sequencing (scRNA-seq) of UC tissue yielded 26 transcriptionally distinct clusters of cells (Figure 2). Based on canonical marker genes, these clusters were annotated to all major cellular components of the bladder tumor

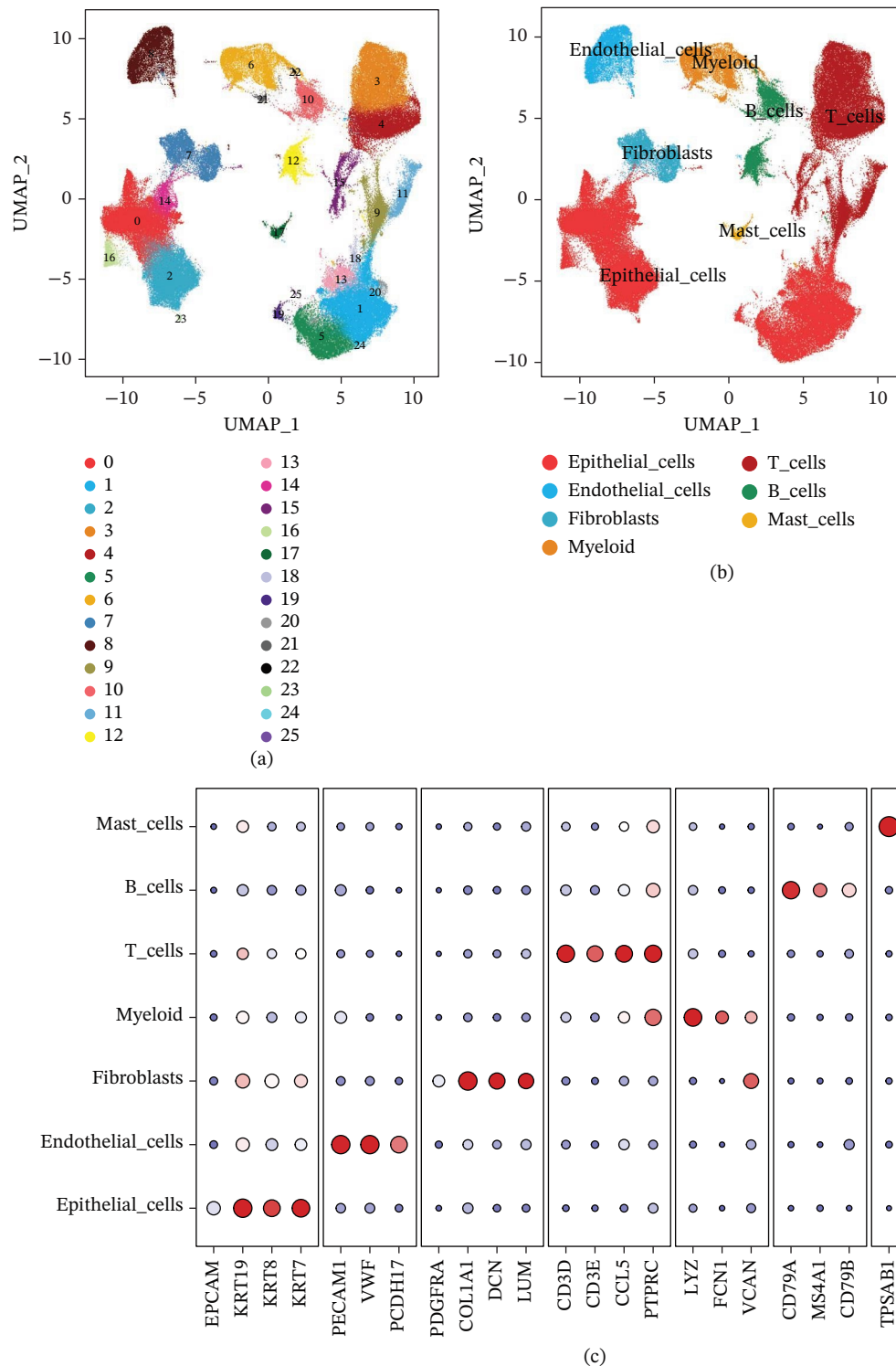


FIGURE 2 | Single-cell transcriptomic analysis of bladder cancer tissue reveals distinct cell populations and their marker genes. (a) UMAP plot of all profiled single cells, colored by unsupervised cluster. (b) UMAP plot of the same cells colored by manually annotated cell type, showing that the clusters correspond to epithelial tumor cells, fibroblasts, endothelial cells, myeloid cells, T cells, B cells, and mast cells. (c) Dot plot of canonical marker gene expression for each cell cluster, with selected marker genes on the x-axis and cell clusters (grouped by cell type) on the y-axis. Dot size represents the percentage of cells in each cluster expressing a given gene, while dot color (color intensity gradient) indicates the average expression level of that gene in the cluster. In the dot plot, each cell type's clusters display the expected marker profile—for example, epithelial cell clusters express epithelial markers (e.g., EPCAM and keratins), fibroblast clusters express stromal markers (COL1A1 for collagen), endothelial clusters express vascular markers (PECAM1/CD31 and VWF), myeloid clusters express myeloid lineage markers (LYZ and CD68), T-cell clusters express T-lymphocyte markers (CD3D and CD3E), B-cell clusters express B lineage markers (CD79A), and the mast cell cluster is characterized by high expression of mast cell tryptase (TPSAB1).

microenvironment, including malignant epithelial cells, fibroblasts, endothelial cells, and diverse immune cells (T cells, B cells, myeloid cells, and a small population of mast cells).

3.6 | Luminal-Like Epithelial Cells Suggesting Immunotherapy Resistance and Enriched in MUT2 Subtype

Upon reclustering, epithelial cells were stratified into 11 distinct clusters. Based on their biological characteristics, three clusters (0, 5, and 7) were classified as neuroendocrine-like epithelial cells (NE1–3), four clusters (1, 3, 4, and 8) as luminal-like epithelial cells (Luminal1–4), and the remaining four clusters (2, 6, 9, and 10) as basal-like epithelial cells (Basal1–4). Notably, luminal-like epithelial cell clusters (Luminal1, Luminal2, and Luminal4) were consistently enriched in the MUT2 subtype across both primary and metastatic UC samples. Furthermore, a lower abundance of the Luminal4 epithelial subpopulation was observed in patients who responded to immunotherapy compared to nonresponders (Figure 3).

3.7 | Fibroblast Subpopulations Are Indicative of Poor Prognosis and Therapeutic Resistance in UC

Through comprehensive transcriptomic profiling, we identified four distinct fibroblast subsets. These included a pericyte-like, contractile CAF cluster (enriched for pericyte markers such as RGS5, NOTCH3, ACTA2, and PDGFRB), a myofibroblastic CAF cluster characterized by high extracellular matrix remodeling gene expression (e.g., CTHRC1, POSTN, and SFRP2), an inflammatory CAF cluster with a pronounced secretory phenotype (expressing cytokines/chemokines like IL6, CXCL12, and CXCL14), and a smooth muscle-like fibroblast subset marked by MYH11 and other smooth muscle genes. Notably, the pericyte-like and myofibroblastic CAFs represent contractile fibroblast phenotypes, whereas the inflammatory CAFs display an immune-regulatory phenotype, highlighting functionally diverse roles within the tumor microenvironment (Figure 4a,b).

To evaluate the clinical significance of these fibroblast subpopulations, we quantified their enrichment in bulk tumors using GSVA scores in TCGA UC cohort. Kaplan–Meier survival analyses demonstrated that patients with high GSVA scores for certain CAF subpopulations had significantly worse OS than those with low scores (Figure 4c–e). For example, tumors enriched in the myofibroblastic signature showed poorer outcomes (HR = 1.39, 95% CI, 1.01–1.91; $p = 0.041$), suggesting that a dense, desmoplastic stroma portends aggressive disease. Likewise, a high inflammatory CAF signature was associated with shortened survival (HR = 1.54, 95% CI, 1.12–2.13; $p = 0.008$), consistent with reports that cytokine-secreting CAFs can suppress antitumor immunity and drive tumor progression. Furthermore, the smooth muscle-like MYH11 fibroblast signature was significantly prognostic (HR = 1.94, 95% CI, 1.4–2.69; $p < 0.001$), indicating a more invasive role for that subpopulation.

We further observed that fibroblast composition differed by tumor subtype and therapy response. The molecular subtype MUT2 exhibited significantly higher GSVA scores for all four

fibroblast subpopulations compared to MUT1 (Figure 4f), implying that MUT2 tumors have a more CAF-rich, stromal microenvironment. In particular, MUT2 tumors (putatively more aggressive) were enriched in inflammatory and pericyte-like CAF signatures relative to MUT1. Finally, analyses of an immunotherapy-treated patient cohort revealed associations between fibroblast subsets and anti-PD-1/PD-L1 treatment response. Patients who responded to immune checkpoint blockade showed a distinct fibroblast signature profile: Notably, responders had diminished inflammatory CAF scores ($p = 0.0062$) and lower pericyte-like and MYH11 CAF scores ($p = 0.01$ and $p = 0.0027$) compared to nonresponders (Figure 4g). Similar results were also observed in endothelial cell subsets (Supporting Information 13: Figure S12). Taken together, these findings highlight how fibroblast heterogeneity in BC impacts tumor biology and clinical outcomes, in line with evidence that activated CAF subsets are linked to advanced disease, immune evasion, and poorer patient survival.

Further analysis of immune-related cells revealed distinct immunological profiles between the MUT1 and MUT2 subtypes. The MUT1 subtype exhibited increased infiltration of plasma cells and plasmablasts, whereas MUT2 demonstrated a higher prevalence of naïve B cells. Similarly, MUT1 showed elevated levels of effector CD8⁺ T cells and cycling CD8⁺ T cells, while MUT2 was characterized by a greater abundance of naïve T cells. This divergence in immune cell infiltration extended to the myeloid compartment, with MUT2 displaying reduced levels of classical CD14⁺ monocytes (Supporting Information 14: Figure S13, Supporting Information 15: Figure S14, and Supporting Information 16: Figure S15). Collectively, these findings suggest a more immunologically active neoplastic milieu in the MUT1 subtype, which may underlie the more favorable clinical outcomes observed in these patients.

3.8 | Machine Learning Performance

Logistic regression achieved the highest predictive performance on the held-out TCGA test set, yielding the greatest area under the ROC curve (AUC). Ensemble methods (Random Forest, LightGBM, XGBoost, CatBoost, and AdaBoost) also achieved high AUC values, slightly below those of logistic regression. In external validation cohorts, logistic regression maintained strong discriminative performance between MUT1 and MUT2, with AUC values generally around or above 0.90 (Figure 5). For example, AUCs reached 0.963 in IMvigor210 and 0.931 in UC-GENOME, representing top-ranking performance in those validations. Performance remained robust even in the largest external set, MSK2022, and the smallest Tongji cohort (AUCs ~0.88 and ~0.84, respectively). Overall, all six machine learning models demonstrated robust and stable performance across cohorts, supporting the robustness and generalizability of the mutation subtype classification.

4 | Discussion

In this study, we developed an approach based on mutational signature CS to identify stable molecular subtypes of UC. This method revealed distinct distributions of underlying

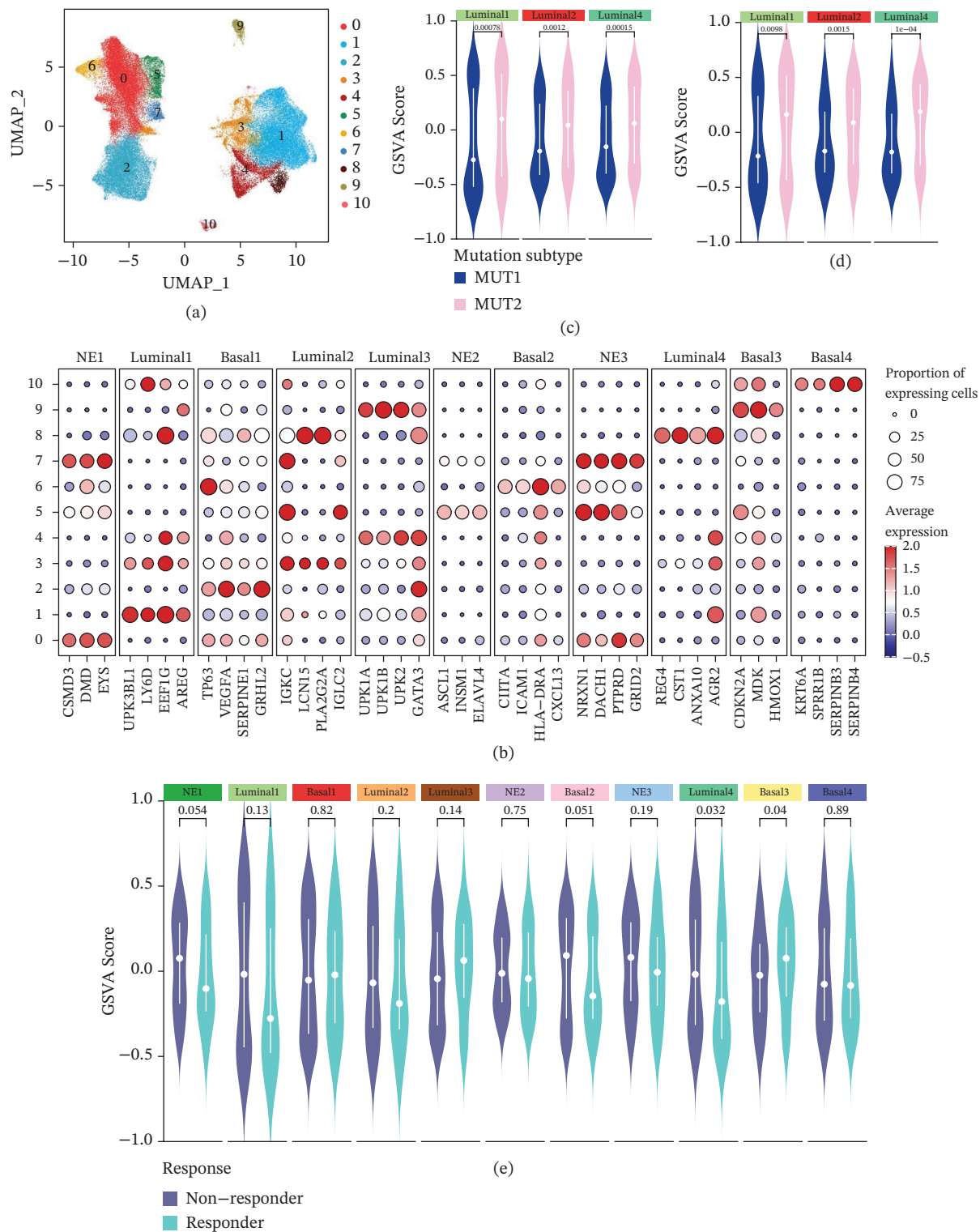


FIGURE 3 | Single-cell-defined epithelial subpopulations and their associations with genomic subtype and immunotherapy response. (a) UMAP embedding of reclustered bladder cancer epithelial cells, colored by 11 transcriptional subpopulations annotated as neuroendocrine-like (NE1–3), luminal-like (Lum1–4), and basal-like (Bas1–4) phenotypes. (b, c) Violin plots of GSVA enrichment scores for luminal-like subtype signatures (Lum1, Lum2, and Lum4) in TCGA and IMvigor210 stratified by mutational subtype (MUT1 vs. MUT2). Each violin shows the score distribution for each subpopulation signature within each mutational subtype; higher scores indicate stronger enrichment of that subtype. Statistical comparisons (two-sided Wilcoxon's rank-sum tests) are indicated. (d) Dot plot of canonical marker gene expression across all 11 epithelial subtypes (columns). (e) Violin plots of GSVA scores for each subtype (NE1–3, Lum1–4, and Bas1–4) comparing patients who responded versus did not respond to immune checkpoint blockade. Each pair of violins shows the score distribution for a given subpopulation in responders versus nonresponders. Statistical significance (Wilcoxon's rank-sum test) is indicated. Abbreviations: GSVA, gene set variation analysis; NE, neuroendocrine-like; Lum, luminal-like; Bas, basal-like.

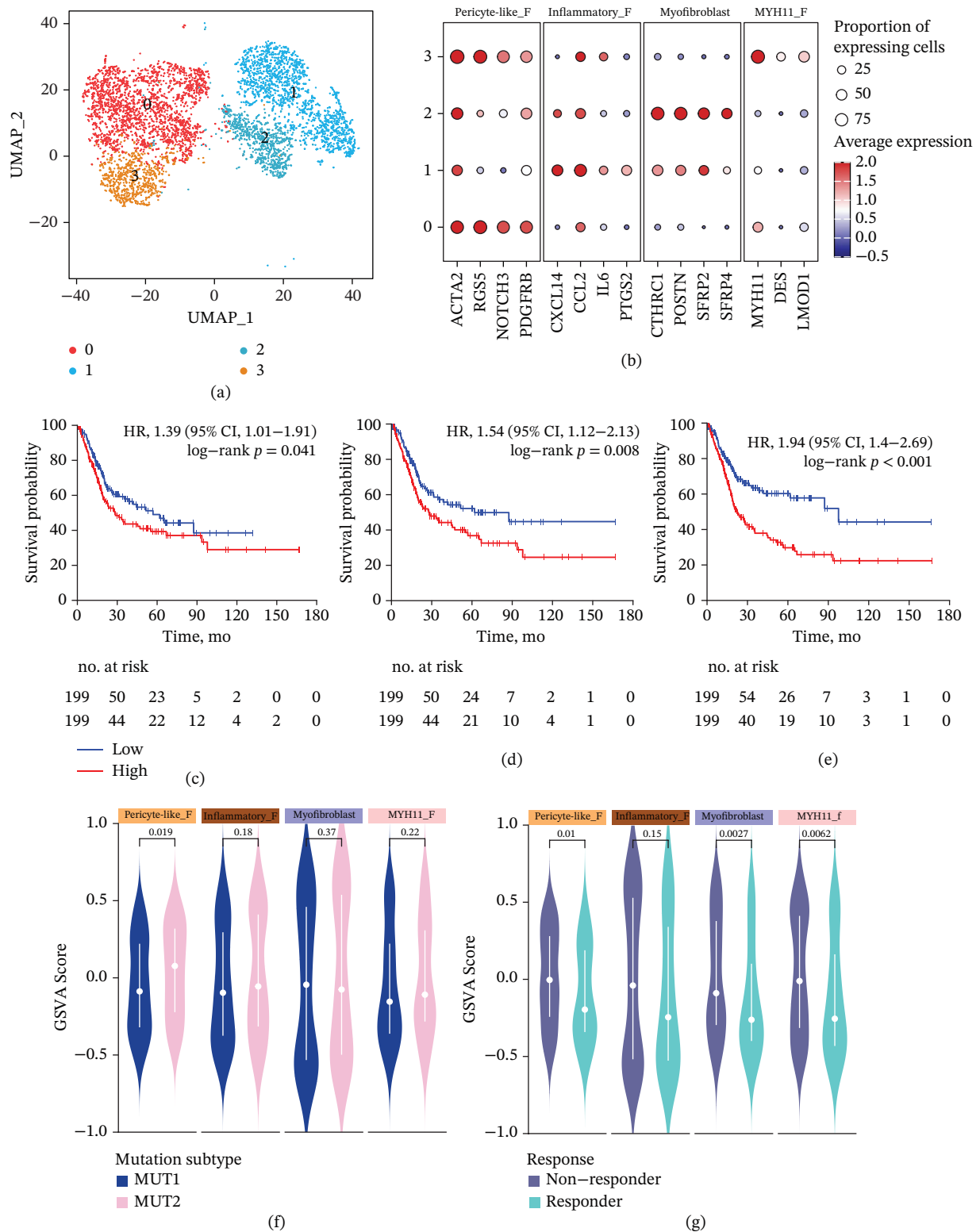


Figure 4 | Legend on next page.

mutational signatures between the MUT1 and MUT2 subtypes and provided biologically meaningful interpretations with corresponding prognostic and therapeutic implications. For example, patients with the MUT1 subtype exhibited enrichment of APOBEC-associated mutational signatures (SBS2 and SBS13). In a mouse melanoma model, APOBEC3B-mediated mutagenesis generated immunogenic neoepitopes in vaccine cells,

activating de novo T-cell responses and enhancing sensitivity to immune checkpoint blockade [58]. Consistent with these findings, our analysis of the IMvigor210 and UC-GENOME cohorts showed that the MUT1 subtype was associated with a higher proportion of tumors with an “inflamed” immune phenotype [58–61]. Furthermore, our mutation classifier showed a strong correlation with the five subtypes of UC defined by Fujii et al.

FIGURE 4 | Fibroblast heterogeneity in bladder cancer and its clinical correlations. (a) UMAP plot of single-cell RNA-seq data from fibroblasts isolated from human bladder tumors. Four distinct fibroblast subtypes are shown. (b) Dot plot of selected canonical marker genes across the four fibroblast subpopulations. (c–e) Kaplan–Meier curves comparing overall survival of bladder cancer patients stratified by high versus low GSVA signature scores for three representative fibroblast subpopulations in TCGA cohort. Hazard ratios (HRs), 95% confidence intervals (CIs), and log-rank *p* values are indicated on each plot, and the number of patients at risk in each group is shown below the x-axis. (f) Violin plots comparing GSVA scores for the four fibroblast subpopulations between two mutational subtypes. Each violin depicts the distribution of signature scores; the white circle denotes the median, the box indicates the interquartile range, and whiskers extend to 1.5× IQR. *p* values above each pair (MUT1 vs. MUT2) were derived from a two-tailed statistical test (Wilcoxon’s rank-sum), showing significant differences in fibroblast signature abundance between the subtypes. (g) Violin plots comparing fibroblast subtype GSVA scores in patients who responded to immune checkpoint inhibitor therapy versus those who did not respond.

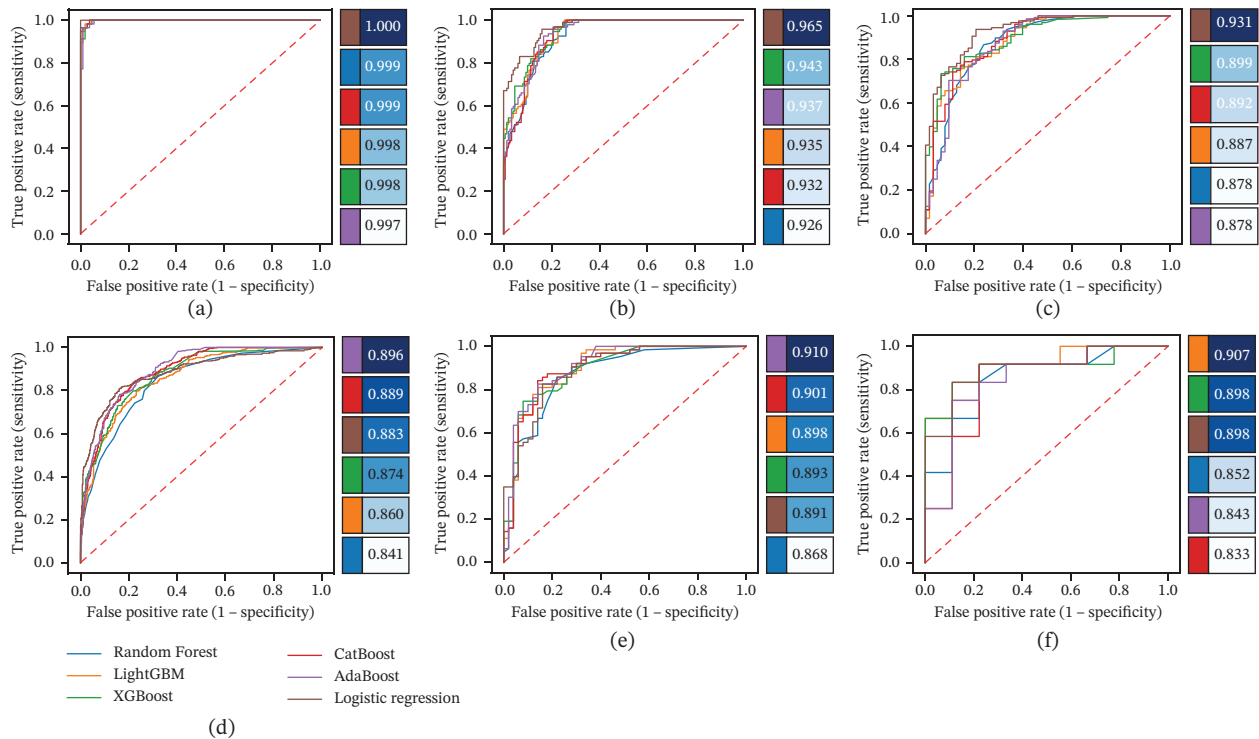


FIGURE 5 | Receiver operating characteristic (ROC) curves comparing the performance of six machine learning classifiers in distinguishing between two mutational subtypes (MUT1 vs. MUT2) across six cohorts. The classifiers include Random Forest, LightGBM, XGBoost, CatBoost, AdaBoost, and logistic regression. All models were trained on TCGA training cohort and evaluated on (a) TCGA internal test set as well as five independent validation cohorts: (b) IMvigor210, (c) UC-GENOME, (d) MSK2022, (e) MSK2015, and (f) Tongji. Each panel corresponds to one dataset and displays the ROC curves for all six models, with the legend indicating the AUC of each model in that cohort.

[26]. Specifically, MUT1 corresponded to the Fujii subtype with a higher proportion of TP53-mutant cases, whereas MUT2 was significantly enriched in the triple-negative subtype. The specific inactivating status of TP53 enhances sensitivity to immunotherapy, not primarily because the tumor becomes “more indolent” but because it becomes “more recognizable to the immune system” under certain molecular contexts. Jin et al. and Lyu et al. collectively point to a central axis involving higher mutational/neoantigen burden, stronger antigen presentation signals, and IFN- γ /CD8-mediated inflammation [62, 63]. Manzano et al. further trace this inflammatory state upstream to features such as a T-cell-inflamed signature, tumor-infiltrating lymphocyte fraction, and APOBEC-related mutational processes [64]. This explains why MUT1 is more sensitive to immunotherapy and is associated with a better prognosis. By contrast, the MUT2 subtype was enriched for “clock-like,” DNA mismatch repair (MMR) deficiency, and exogenous chemotherapy-associated mutational signatures (e.g., SBS1, SBS6, SBS15, and SBS87). This pattern aligns with our Cox regression analysis, which

identified older age as a risk factor for UC, and with subgroup analyses showing that the MUT2 subtype was more prevalent in older patients. Both clinical evidence and cell/animal models suggest that MMR deficiency (reflected by SBS6/15 signatures) is associated with resistance to platinum-based chemotherapy [65, 66]. Loss of MMR function impairs the ability of tumor cells to initiate apoptosis in response to platinum-induced DNA damage, thereby promoting the selection of resistant clones and reducing treatment efficacy. Similar mechanisms have been reported in other malignancies, including MMR-deficient lung and ovarian cancers, suggesting that comparable resistance mechanisms may operate in MMR-deficient UC.

Our findings indicate that patients with the MUT2 subtype, characterized by MMR deficiency-associated signatures, have a poorer prognosis than those with the MUT1 subtype, which is enriched for APOBEC signatures. This prognostic difference was observed across primary and metastatic settings and was consistent across multiple clinical subgroups. Specifically,

MUT2 tumors exhibit a higher abundance of CAFs, including inflammatory and pericyte-like CAF subpopulations, along with reduced infiltration of immune effector cells associated with reduced responsiveness to immunotherapy. We observed this trend among patients with a history of smoking, those with an ECOG performance status of 1, those lacking PD-L1 expression on immune (IC0) and tumor cells (TC0), and those with the “desert” immunophenotype. A similar risk differential was noted even among patients who received platinum-based chemotherapy and achieved a clinical benefit; in each of these contexts, the MUT2 subtype was associated with a significantly higher risk of death than the MUT1 subtype. These differences in prognostic and therapeutic response between patients with the MUT2 and MUT1 subtypes may reflect alterations in the tumor microenvironment driven by their distinct genomic profiles. Our findings offer insights for clinical practice by providing more refined mortality risk assessments and guiding the development of precise individualized treatment strategies.

This study has several limitations. First, differences in exome sequencing approaches across cohorts may have influenced the detection of mutational signatures because targeted sequencing panels typically cover fewer genes (and thus capture fewer mutations) than WES. Second, CS (a mathematical measure of how closely the shape of a sample’s mutational pattern matches a reference signature) was used to quantify the similarity between tumor mutational profiles and reference signatures; moderate similarity may reflect a mixture of multiple mutational processes rather than a single dominant signature. Third, although the machine learning models demonstrated strong performance across retrospective datasets, prospective validation and calibration will be required to determine optimal thresholds and implementation in clinical settings.

Despite these limitations, our findings were consistent across multiple cohorts and sequencing platforms, supporting the robustness and generalizability of this approach. Classification based on mutational signature similarity enabled reproducible risk stratification and provided insights into the biological mechanisms underlying subtype-specific differences in treatment response. By linking mutational patterns to underlying carcinogenic processes, this framework may also facilitate the identification of novel therapeutic targets and improve patient stratification strategies.

5 | Conclusion

In this retrospective study, we identified two distinct mutational signature subtypes of UC (MUT1 and MUT2), reflecting underlying tumor heterogeneity. MUT1, characterized predominantly by APOBEC-associated mutational signatures, was associated with a more favorable prognosis and higher immunotherapy response rates. In contrast, MUT2, characterized by MMR deficiency-related signatures, was associated with poorer outcomes, including increased risk of death, particularly among patients receiving chemotherapy, with a history of smoking, or displaying certain immune phenotypes. These findings provide novel insights into clinical decision-making for personalized treatment strategies and will help improve patient outcomes in UC by facilitating an exploration of resistance mechanisms,

immunotherapy optimization, and long-term outcomes and survivorship.

Author Contributions

Chaozhi Tang was responsible for conceptualization, methodology, data curation, investigation, validation, formal analysis, visualization, and writing of the original draft, as well as review and editing of the manuscript. Yifan Liu and Tian-Long Wang were responsible for methodology, software, data curation, validation, writing of the original draft, and resources. Jiakang Ma was responsible for methodology and software. Weihua Chen was responsible for data curation and investigation. Xiang Liu, Jingdong Xue, Lin Ye, and Feng Yue were responsible for methodology, resources, and supervision. Chaozhi Tang, Yifan Liu, and Tian-Long Wang have contributed to the work equally and should be regarded as cofirst authors.

Acknowledgments

This study was made possible by the generous sharing of the IMvigor210 whole-exome sequencing dataset (EGAS00001002556) by Mariathasan et al., for which we are deeply grateful. We further acknowledge the insightful theoretical counsel provided by Dr. Jeffrey S. Damrauer. Additionally, we thank Editage for their professional language polishing services for the graphical abstract. During the preparation of this work, the authors used ChatGPT (GPT-4o) in order to improve readability and clarity. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of this publication.

Funding

This work was supported by the Shanghai Xuhui Medical Research Project (Grant Numbers SHXH202336 and SHXH202513), the Shanghai Eighth People’s Hospital Research Project (Grant Number SHBY202513), and the Medical Scientific Research Foundation of Guangdong Province, China (Grant Number B2022055).

Disclosure

All authors have read and approved the final manuscript.

Ethics Statement

This study was reviewed and approved by the Ethical Committee of the Shanghai East Hospital (No. 2018YS-04) and complied with the principles of the Declaration of Helsinki. Written informed consent was obtained from all patients.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding authors upon reasonable request.

References

1. F. Bray, M. Laversanne, H. Sung, et al., “Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *Ca: Cancer Journal for Clinicians* 74, no. 3 (2024): 229–263, <https://doi.org/10.3322/caac.21834>.
2. I. Jubber, S. Ong, L. Bukavina, et al., “Epidemiology of Bladder Cancer in 2023: A Systematic Review of Risk Factors,” *European Urology* 84, no. 2 (2023): 176–190, <https://doi.org/10.1016/j.eururo.2023.03.029>.

3. C. Svendsen, A. James, R. S. Matulewicz, et al., "Carcinogenic Biomarkers of Exposure in the Urine of Heated Tobacco Product Users Associated With Bladder Cancer: A Systematic Review," *Urologic Oncology* 40, no. 4 (2022): 149–160, <https://doi.org/10.1016/j.urolonc.2021.11.018>.
4. N. Lobo, L. Afferi, M. Moschini, et al., "Epidemiology, Screening, and Prevention of Bladder Cancer," *European Urology Oncology* 5, no. 6 (2022): 628–639, <https://doi.org/10.1016/j.euo.2022.10.003>.
5. A. T. Lenis, P. M. Lec, K. Chamie, and M. D. Mshs, "Bladder Cancer," *Jama* 324, no. 19 (2020): 1980–1991, <https://doi.org/10.1001/jama.2020.17598>.
6. A. G. van der Heijden, H. M. Bruins, A. Carrion, et al., "European Association of Urology Guidelines on Muscle-Invasive and Metastatic Bladder Cancer: Summary of the 2025 Guidelines," *European Urology* 87, no. 5 (2025): 582–600, <https://doi.org/10.1016/j.eururo.2025.02.019>.
7. H. Su, H. Jiang, T. Tao, et al., "Hope and Challenge: Precision Medicine in Bladder Cancer," *Cancer Medicine* 8, no. 4 (2019): 1806–1816, <https://doi.org/10.1002/cam4.1979>.
8. F. Audenet, K. Attalla, and J. P. Sfakianos, "The Evolution of Bladder Cancer Genomics: What Have We Learned and How Can We Use It?," *Urologic Oncology* 36, no. 7 (2018): 313–320, <https://doi.org/10.1016/j.urolonc.2018.02.017>.
9. M. Ferro, F. Crocetto, L. M. I. Jannello, et al., "Emerging Biomarkers in Bladder Cancer: Translating Molecular Advances Into Precision Oncology," *Critical Reviews in Oncology Hematology* 215 (2025): 104862, <https://doi.org/10.1016/j.critrevonc.2025.104862>.
10. S. K. Mohanty, A. Lobo, S. K. Mishra, and L. Cheng, "Precision Medicine in Bladder Cancer: Present Challenges and Future Directions," *Journal of Personalized Medicine* 13, no. 5 (2023): 756, <https://doi.org/10.3390/jpm13050756>.
11. L. Tran, J. F. Xiao, N. Agarwal, J. E. Duex, and D. Theodorescu, "Advances in Bladder Cancer Biology and Therapy," *Nature Reviews Cancer* 21, no. 2 (2021): 104–121, <https://doi.org/10.1038/s41568-020-00313-1>.
12. W. Choi, S. Porten, S. Kim, et al., "Identification of Distinct Basal and Luminal Subtypes of Muscle-Invasive Bladder Cancer With Different Sensitivities to Frontline Chemotherapy," *Cancer Cell* 25, no. 2 (2014): 152–165, <https://doi.org/10.1016/j.ccr.2014.01.009>.
13. A. Kamoun, A. de Reynies, Y. Allory, et al., "A Consensus Molecular Classification of Muscle-Invasive Bladder Cancer," *European Urology* 77, no. 4 (2020): 420–433, <https://doi.org/10.1016/j.eururo.2019.09.006>.
14. K. Jin, Y. Ding, J. Xu, et al., "Lethal Clinical Outcome and Chemotherapy and Immunotherapy Resistance in Patients With Urothelial Carcinoma With MDM2 Amplification or Overexpression," *Journal for ImmunoTherapy of Cancer* 13, no. 1 (2025): , <https://doi.org/10.1136/jitc-2024-010964> e010964.
15. A. T. Lenis, K. Whiting, V. Ravichandran, et al., "Natural History and Genomic Landscape of Chemotherapy-Resistant Muscle-Invasive Bladder Cancer," *JCO Precision Oncology* 8, no. 8 (2024): e2300274, <https://doi.org/10.1200/PO.23.00274>.
16. B. J. Guercio, M. Sarfaty, M. Y. Teo, et al., "Clinical and Genomic Landscape of FGFR3-Altered Urothelial Carcinoma and Treatment Outcomes With Erdafitinib: A Real-World Experience," *Clinical Cancer Research* 29, no. 22 (2023): 4586–4595, <https://doi.org/10.1158/1078-0432.CCR-23-1283>.
17. F. Facchinetti, A. Hollebecque, F. Braye, et al., "Resistance to Selective FGFR Inhibitors in FGFR-Driven Urothelial Cancer," *Cancer Discovery* 13, no. 9 (2023): 1998–2011, <https://doi.org/10.1158/2159-8290.CD-22-1441>.
18. M. Y. Teo, J. M. Mota, K. A. Whiting, et al., "Fibroblast Growth Factor Receptor 3 Alteration Status Is Associated With Differential Sensitivity to Platinum-Based Chemotherapy in Locally Advanced and Metastatic Urothelial Carcinoma," *European Urology* 78, no. 6 (2020): 907–915, <https://doi.org/10.1016/j.eururo.2020.07.018>.
19. B. Miron, J. H. Hoffman-Censits, F. Anari, et al., "Defects in DNA Repair Genes Confer Improved Long-term Survival After Cisplatin-Based Neoadjuvant Chemotherapy for Muscle-Invasive Bladder Cancer," *European Urology Oncology* 3, no. 4 (2020): 544–547, <https://doi.org/10.1016/j.euo.2020.02.003>.
20. V. McPherson, B. Reardon, A. Bhayankara, et al., "A Phase 2 Trial of Buparlisib in Patients With Platinum-Resistant Metastatic Urothelial Carcinoma," *Cancer* 126, no. 20 (2020): 4532–4544, <https://doi.org/10.1002/cncr.33071>.
21. G. Iyer, A. V. Balar, M. I. Milowsky, et al., "Multicenter Prospective Phase II Trial of Neoadjuvant Dose-Dense Gemcitabine Plus Cisplatin in Patients With Muscle-Invasive Bladder Cancer," *Journal of Clinical Oncology* 36, no. 19 (2018): 1949–1956, <https://doi.org/10.1200/JCO.2017.75.0158>.
22. D. Liu, E. R. Plimack, J. Hoffman-Censits, et al., "Clinical Validation of Chemotherapy Response Biomarker ERCC2 in Muscle-Invasive Urothelial Bladder Carcinoma," *JAMA Oncology* 2, no. 8 (2016): 1094–1096, <https://doi.org/10.1001/jamaoncol.2016.1056>.
23. E. R. Plimack, R. L. Dunbrack, T. A. Brennan, et al., "Defects in DNA Repair Genes Predict Response to Neoadjuvant Cisplatin-Based Chemotherapy in Muscle-Invasive Bladder Cancer," *European Urology* 68, no. 6 (2015): 959–967, <https://doi.org/10.1016/j.eururo.2015.07.009>.
24. E. M. Van Allen, K. W. Mouw, P. Kim, et al., "Somatic ERCC2 Mutations Correlate With Cisplatin Sensitivity in Muscle-Invasive Urothelial Carcinoma," *Cancer Discovery* 4, no. 10 (2014): 1140–1153, <https://doi.org/10.1158/2159-8290.CD-14-0623>.
25. E. Blaveri, J. P. Simko, J. E. Korkola, et al., "Bladder Cancer Outcome and Subtype Classification by Gene Expression," *Clinical Cancer Research* 11, no. 11 (2005): 4044–4055, <https://doi.org/10.1158/1078-0432.CCR-04-2409>.
26. Y. Fujii, Y. Sato, H. Suzuki, et al., "Molecular Classification and Diagnostics of Upper Urinary Tract Urothelial Carcinoma," *Cancer Cell* 39, no. 6 (2021): 793–809, <https://doi.org/10.1016/j.ccell.2021.05.008>.
27. F. A. Sinicrope, K. Okamoto, P. M. Kasi, and H. Kawakami, "Molecular Biomarkers in the Personalized Treatment of Colorectal Cancer," *Clinical Gastroenterology and Hepatology* 14, no. 5 (2016): 651–658, <https://doi.org/10.1016/j.cgh.2016.02.008>.
28. J. A. Nakauma-Gonzalez, M. Rijnders, J. van Riet, et al., "Comprehensive Molecular Characterization Reveals Genomic and Transcriptomic Subtypes of Metastatic Urothelial Carcinoma," *European Urology* 81, no. 4 (2022): 331–336, <https://doi.org/10.1016/j.eururo.2022.01.026>.
29. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, et al., "Signatures of Mutational Processes in Human Cancer," *Nature* 500, no. 7463 (2013): 415–421, <https://doi.org/10.1038/nature12477>.
30. J. Le, Y. Dian, D. Zhao, et al., "Single-Cell Multi-Omics in Cancer Immunotherapy: From Tumor Heterogeneity to Personalized Precision Treatment," *Molecular Cancer* 24, no. 1 (2025): 221, <https://doi.org/10.1186/s12943-025-02426-3>.
31. J. Chen, L. Larsson, A. Swarbrick, and J. Lundeberg, "Spatial Landscapes of Cancers: Insights and Opportunities," *Nature Reviews Clinical Oncology* 21, no. 9 (2024): 660–674, <https://doi.org/10.1038/s41571-024-00926-7>.
32. S. Kim, G. Leem, J. Choi, et al., "Integrative Analysis of Spatial and Single-Cell Transcriptome Data From Human Pancreatic Cancer Reveals an Intermediate Cancer Cell Population Associated With Poor Prognosis," *Genome Medicine* 16, no. 1 (2024): 20, <https://doi.org/10.1186/s13073-024-01287-7>.
33. K. Fang, A. G. Ohiohoin, T. Liu, et al., "Integrated Single-Cell Analysis Reveals Distinct Epigenetic-Regulated Cancer Cell States and

- a Heterogeneity-Guided Core Signature in Tamoxifen-Resistant Breast Cancer,” *Genome Medicine* 16, no. 1 (2024): 134, <https://doi.org/10.1186/s13073-024-01407-3>.
34. Y. Luo, and H. Liang, “Single-Cell Dissection of Tumor Microenvironmental Response and Resistance to Cancer Therapy,” *Trends in Genetics* 39, no. 10 (2023): 758–772, <https://doi.org/10.1016/j.tig.2023.07.005>.
35. O. Elhanani, R. Ben-Uri, and L. Keren, “Spatial Profiling Technologies Illuminate the Tumor Microenvironment,” *Cancer Cell* 41, no. 3 (2023): 404–420, <https://doi.org/10.1016/j.ccell.2023.01.010>.
36. A. G. Robertson, J. Kim, H. Al-Ahmadie, et al., “Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer,” *Cell* 171, no. 3 (2017): 540–556.e25, <https://doi.org/10.1016/j.cell.2017.09.007>.
37. S. Mariathasan, S. J. Turley, D. Nickles, et al., “TGF β Attenuates Tumour Response to PD-L1 Blockade by Contributing to Exclusion of T Cells,” *Nature* 554, no. 7693 (2018): 544–548, <https://doi.org/10.1038/nature25501>.
38. J. E. Rosenberg, J. Hoffman-Censits, T. Powles, et al., “Atezolizumab in Patients With Locally Advanced and Metastatic Urothelial Carcinoma Who Have Progressed Following Treatment With Platinum-Based Chemotherapy: A Single-Arm, Multicentre, Phase 2 Trial,” *Lancet* 387, no. 10031 (2016): 1909–1920, [https://doi.org/10.1016/S0140-6736\(16\)00561-4](https://doi.org/10.1016/S0140-6736(16)00561-4).
39. J. S. Damrauer, W. Beckabir, J. Klomp, et al., “Collaborative Study From the Bladder Cancer Advocacy Network for the Genomic Analysis of Metastatic Urothelial Cancer,” *Nature Communications* 13, no. 1 (2022): 6658, <https://doi.org/10.1038/s41467-022-33980-9>.
40. J. P. Sfakianos, E. K. Cha, G. Iyer, et al., “Genomic Characterization of Upper Tract Urothelial Carcinoma,” *European Urology* 68, no. 6 (2015): 970–977, <https://doi.org/10.1016/j.eururo.2015.07.039>.
41. T. N. Clinton, Z. Chen, H. Wise, et al., “Genomic Heterogeneity as a Barrier to Precision Oncology in Urothelial Cancer,” *Cell Reports* 41, no. 12 (2022): 111859, <https://doi.org/10.1016/j.celrep.2022.111859>.
42. A. M. Bolger, M. Lohse, and B. Usadel, “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data,” *Bioinformatics* 30, no. 15 (2014): 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170>.
43. H. Li, and R. Durbin, “Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform,” *Bioinformatics* 25, no. 14 (2009): 1754–1760, <https://doi.org/10.1093/bioinformatics/btp324>.
44. H. Li, B. Handsaker, A. Wysoker, et al., “The Sequence Alignment/Map Format and SAMtools,” *Bioinformatics* 25, no. 16 (2009): 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
45. K. Wang, M. Li, and H. Hakonarson, “ANNOVAR: Functional Annotation of Genetic Variants From High-Throughput Sequencing Data,” *Nucleic Acids Research* 38, no. 16 (2010): e164, <https://doi.org/10.1093/nar/gkq603>.
46. A. McKenna, M. Hanna, E. Banks, et al., “The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data,” *Genome Research* 20, no. 9 (2010): 1297–1303, <https://doi.org/10.1101/gr.107524.110>.
47. W. McLaren, L. Gil, S. E. Hunt, et al., “The Ensembl Variant Effect Predictor,” *Genome Biology* 17, no. 1 (2016): 122, <https://doi.org/10.1186/s13059-016-0974-4>.
48. J. S. Gehring, B. Fischer, M. Lawrence, and W. Huber, “Somatic Signatures: Inferring Mutational Signatures From Single-Nucleotide Variants,” *Bioinformatics* 31, no. 22 (2015): 3673–3675, <https://doi.org/10.1093/bioinformatics/btv408>.
49. J. G. Tate, S. Bamford, H. C. Jubb, et al., “COSMIC: The Catalogue of Somatic Mutations in Cancer,” *Nucleic Acids Research* 47, no. D1 (2019): D941–D947, <https://doi.org/10.1093/nar/gky1015>.
50. T. Stuart, A. Butler, P. Hoffman, et al., “Comprehensive Integration of Single-Cell Data,” *Cell* 177, no. 7 (2019): 1888–1902.e21, <https://doi.org/10.1016/j.cell.2019.05.031>.
51. S. Hanzelmann, R. Castelo, and J. Guinney, “GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data,” *BMC Bioinformatics* 14, no. 1 (2013): 7, <https://doi.org/10.1186/1471-2105-14-7>.
52. L. Breiman, “Random Forests,” *Machine Learning* 45, no. 1 (2001): 5–32, <https://doi.org/10.1023/A:1010933404324>.
53. G. Ke, Q. Meng, T. Finley, et al., “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” *Advances in Neural Information Processing Systems* 30 (2017): 3149–3157.
54. T. Chen, and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Digital Library, 2016), 785–794, <https://doi.org/10.1145/2939672.2939785>.
55. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: Unbiased Boosting With Categorical Features,” in *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (Curran Associates, Inc, 2018), 6639–6649.
56. Y. Freund, and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” *Journal of Computer and System Sciences* 55, no. 1 (1997): 119–139, <https://doi.org/10.1006/jcss.1997.1504>.
57. D. R. Cox, “The Regression Analysis of Binary Sequences,” *Journal of the Royal Statistical Society: Series B (Methodological)* 21, no. 1 (1959): 238–238, <https://doi.org/10.1111/j.2517-6161.1959.tb00334.x>.
58. C. B. Driscoll, M. R. Schuelke, T. Kottke, et al., “APOBEC3B-Mediated Corruption of the Tumor Cell Immunopeptidome Induces Heteroclitic Neoepitopes for Cancer Immunotherapy,” *Nature Communications* 11, no. 1 (2020): 790, <https://doi.org/10.1038/s41467-020-14568-7>.
59. A. Gupta, A. Gazzo, P. Selenica, et al., “APOBEC3 Mutagenesis Drives Therapy Resistance in Breast Cancer,” *Nature Genetics* 57, no. 6 (2025): 1452–1462, <https://doi.org/10.1038/s41588-025-02187-1>.
60. X. Li, Y. Wang, S. Deng, et al., “Loss of SYNCYRIP Unleashes APOBEC-Driven Mutagenesis, Tumor Heterogeneity, and AR-Targeted Therapy Resistance in Prostate Cancer,” *Cancer Cell* 41, no. 8 (2023): 1427–1449.e12, <https://doi.org/10.1016/j.ccell.2023.06.010>.
61. E. K. Law, A. M. Sieuwerts, K. LaPara, et al., “The DNA Cytosine Deaminase APOBEC3B Promotes Tamoxifen Resistance in ER-Positive Breast Cancer,” *Science Adventure* 2, no. 10 (2016): e1601737, <https://doi.org/10.1126/sciadv.1601737>.
62. K. Jin, J. Xu, X. Su, et al., “TP53 Disruptive Mutation Predicts Platinum-Based Chemotherapy and PD-1/PD-L1 Blockade Response in Urothelial Carcinoma,” *Journal of Pathology* 263, no. 2 (2024): 139–149, <https://doi.org/10.1002/path.6266>.
63. Q. Lyu, A. Lin, M. Cao, A. Xu, P. Luo, and J. Zhang, “Alterations in TP53 Are a Potential Biomarker of Bladder Cancer Patients Who Benefit From Immune Checkpoint Inhibition,” *Cancer Control* 27, no. 1 (2020): 1073274820976665, <https://doi.org/10.1177/1073274820976665>.
64. R. G. Manzano, A. Catalan-Latorre, and A. Brugarolas, “RB1 and TP53 Co-Mutations Correlate Strongly With Genomic Biomarkers of Response to Immunity Checkpoint Inhibitors in Urothelial Bladder Cancer,” *BMC Cancer* 21, no. 1 (2021): 432, <https://doi.org/10.1186/s12885-021-08078-y>.
65. H. S. Hsu, C. K. Wen, Y. A. Tang, et al., “Promoter Hypermethylation Is the Predominant Mechanism in hMLH1 and hMSH2 Deregulation and Is a Poor Prognostic Factor in Nonsmoking Lung Cancer,” *Clinical Cancer Research* 11, no. 15 (2005): 5410–5416, <https://doi.org/10.1158/1078-0432.CCR-05-0601>.
66. G. Gifford, J. Paul, P. A. Vasey, S. B. Kaye, R. Brown, and on behalf of the Scottish Gynaecological Clinical Trials Group, “The Acquisition of hMLH1 Methylation in Plasma DNA After Chemotherapy Predicts Poor

Supporting Information

Additional supporting information can be found online in the Supporting Information 1. **Supporting Information.** Methods S1: Computation and interpretation of cosine similarity. Methods S2: Identification of the mutational subtype of urothelial carcinoma. Methods S3: Identification of the risk score for urothelial carcinoma. Methods S4: Development and external validation of a machine learning-based prediction model. Supporting Information 2. Figure S1: Non-negative matrix factorization (NMF) of the cosine similarity matrix of mutation signatures from TCGA cohort. Supporting Information 3. Figure S2: Identification and clinical characteristics of mutational signature classification in TCGA training cohort. Supporting Information 4. Figure S3: Clinical characteristics of mutational signature classifications in the MSK2022 test cohort. Supporting Information 5. Figure S4: Clinical characteristics according to the mutational signature classification in the MSK2015 test cohort. Supporting Information 6. Figure S5: Clinical characteristics of mutational signature classification in the IMvigor210 and UC-GENOME test cohorts. Supporting Information 7. Figure S6: Kaplan–Meier curves depicting overall survival (OS) in patients stratified by the genomic mutation-based risk score: $\text{score} < 1$ (low risk) versus ≥ 1 (high risk). Supporting Information 8. Figure S7: Differences in clinical characteristics according to the mutation signature subtype and immunotherapy response. Supporting Information 9. Figure S8: Subgroup analysis of clinical characteristics based on the mutational subtype (IMvigor210 cohort). Supporting Information 10. Figure S9: Subgroup analysis of clinical characteristics based on the mutational subtype (UC-GENOME cohort). Supporting Information 11. Figure S10: Univariate analysis for mutational signature subtype and clinical characteristics across multiple cohorts. Supporting Information 12. Figure S11: Multivariate analysis for mutational signature subtype and clinical characteristics across multiple cohorts. Supporting Information 13. Figure S12: Endothelial subpopulations are indicative of poor prognosis and therapeutic resistance in urothelial carcinoma. Supporting Information 14. Figure S13: Distribution of myeloid cells in mutational subtypes. Supporting Information 15. Figure S14: Distribution of B cells in mutational subtypes. Supporting Information 16. Figure S15: Distribution of T cells in mutational subtypes. Supporting Information 17. Table S1: Mutational signatures used in mutational subtype classifier. Table S2: Association of genetic alterations with overall survival in the multivariable regression analysis. Table S3: Distribution of clinical and genomic features by mutational signature subtype in the IMvigor210 cohort. Table S4: Distribution of clinical and genomic features by mutational signature subtype in the UC-GENOME cohort. Table S5: Distribution of clinical and genomic features by mutational signature subtype in TCGA cohort. Table S6: Distribution of clinical and genomic features by mutational signature subtype in the MSK2022 cohort. Table S7: Distribution of clinical and genomic features by mutational signature subtype in the MSK2015 cohort.